

“华为杯”第十五届中国研究生 数学建模竞赛

题 目 光传送网建模与价值评估

摘 要：

本文针对光传送网的链路，利用二维正态分布的概率模型对不同调制格式的纠前错误概率进行了估计，并建立了光传输链路的传输性能模型。针对光网络的规划问题，建立了赋权连通图模型，利用 Prim 算法、最速下降法、Floyd 算法以及模拟退火算法进行优化，进而得到近似全局最优解。针对不等概率的调制方式提出了一种优化的星座图配置。

针对问题一(1)，我们探究 QPSK、8QAM 和 16QAM 三种调制格式，在不同信噪比 SNR 情况下接收到的调制信号概率分布。利用二维正态分布对接收到的符号的概率分布进行了建模，在信号功率归一化的约束条件下，对每一类星座图上的星座点的出现概率以及判决错误概率期望进行了估计，结合每个星座点的信息熵计算出了纠前错误概率和 SNR 的关系曲线。通过二分法得出了 $BER = 0.02$ 时的 SNR 容限点，对 QPSK、8QAM 和 16QAM 三种调制方式 6.9207、10.341 和 12.7726。问题一(2)，要求计算光链路在符合 SNR 容限情况下最大传播跨数和最远传播距离。我们建立了光链路传输过程中的信号噪声功率衰减、增益的动态模型，求解出了在跨长为 80km 时，QPSK、8QAM 和 16QAM 能传输 17, 8, 4 跨，其最远传输距离为 1360, 640, 320km；在跨长为 100km 时，

QPSK、8QAM 和 16QAM 能传输 6, 3, 1 跨, 其最远传输距离为 600, 300, 100km;

针对问题二(1), 我们将光传送网规划简化成赋权连通图问题, 通过 prim 算法求得 12 个城市的最大生成树, 并将剩下的最大价值边加入到传送网里, 16 条连接和 33 条连接的最大价值分别为 5174.3mTb/s 和 8418.6mTb/s。针对问题二(2), 我们结合数学推导和最速下降法编程证明, 定性、定量的证明了:

- 1) 城市互连结构里, 即使考虑分流机制, 也是最大生成树价值最大;
- 2) 添加成树之外的边时三元结构中引入分流机制所得价值比当前最大价值要低;
- 3) 证明添加成树之外的边时任意元结构中引入分流机制所得价值比当前最大价值要低。

从而在最大生成树基础上, 综合考虑城市两两间最低通信容量 $\min bw = 0.1 \text{ TB/s}$, 得到引入中转分流节点下 16 条连接和 33 条连接的城市最大网络价值分别为 7069.9mTb/s 和 19151.46mTb/s, 16 条连接和 33 条连接的省份最大网络价值分别为 23510.5mTb/s 和 64254.4mTb/s。针对问题二(3), 我们综合考虑了城市距离、人口、发达程度、辐射范围以及地理条件, 既考虑了点单独的信息, 又考虑了点与点之间的邻域信息, 提出了新的目标价值函数 $E_{ij} =$

$\mu CP_{ij}^{\alpha} G_{ij}^{\beta} (D_{max} - D_{ij})^{\gamma}$ 。并根据此目标函数结合 Floyd 算法以及模拟退火算法, 给出了 16 条连接和 33 条连接的最优网络连接图, 使得华南地区以武汉和上海为中心, 华北地区以北京为中心, 并且对于西南地区以及新疆的规划也都更加合理, 同时也保持连通所必要的跨地形阶梯连接。

针对问题三, 我们探寻满足信息熵为 3bit 的 16QAM 的基础上优化星座图上星座点的位置分布以及出现概率, 以达到性能的最优。根据对问题的分析, 我 9QAM, 利用对二的负幂次方概率的穷举式搜索方式, 确定了一组符合条件星座图概率分布 $(\frac{1}{4}, \frac{1}{8} \times 4, \frac{1}{16} \times 4)$, 结合格雷码编码优化出了一种非等概率的 9QAM 的调制方式。其 SNR 容限点为 8.7321dB, 较题设的 8QAM 调制方式优化了 1.6 dB。

关键词: 光传送网规划 二维高斯分布 最大生成树 Floyd 算法 模拟退火算法 QAM 调制

1 问题重述

1.1 问题背景

2009 年诺贝尔物理学奖授予了英籍华人高锟（Charles K. Kao）博士，以表彰他对光纤通信发展所做出的贡献，诺贝尔奖委员会在给公众的公开信中写到：

“当诺贝尔物理学奖宣布的时候，世界大部分地方几乎瞬间收到了这条信息…文字、语音和视频信号沿着光纤在世界各地来回传输，几乎瞬时地被微小而便捷的设备接收，人们已经把这种情况当做习惯。光纤通信正是整个通信领域急速发展的前提。”

从诞生至今，50 多年里基于数字光纤通信技术的光传送网构建起了全球通信的骨架。从城市内的传输，直到跨越大洋的传输，光传送网为人类提供了大容量、高可靠性和低能耗的信息传输管道，人类对通信容量的追求也成为光传送技术发展的源源不断的动力。

光传送网的规划与建设是运营商、设备商以及政府必须考虑的课题。光传送的基本规律是——在相同技术条件下传输的容量会随着传输距离增加而减小。网络规划者需要在有限资源的条件下，综合考虑传输距离，传输容量、网络拓扑等各种因素，以最大化网络的价值。本课题中，请你们站在上述角度，从底层物理出发为光传送链路建模，制定光传送网规划，探索光传送网有关规律。

1.2 需要解决的问题

1.2.1 问题一：光传送链路建模

现代数字传输系统可认为是对 0101 二进制序列进行编码传输的系统，1 个二进制的 0 或 1 称为 1 个比特（bit）。无论是语音、视频还是任何类型的消息，都可以数字化为一串串“0101...”的二进制比特序列，经编码并调制为某个“载体信号”后，再经过特定的“信道”（信息的通道）传输到目的地。图 1 中给出了简化的模型。在光纤通信中，光纤就是信道，光纤传输的光波就是信息的载体。信道中无法避免的噪声可能导致最终接收的二进制序列中比特出错，即产生误码。

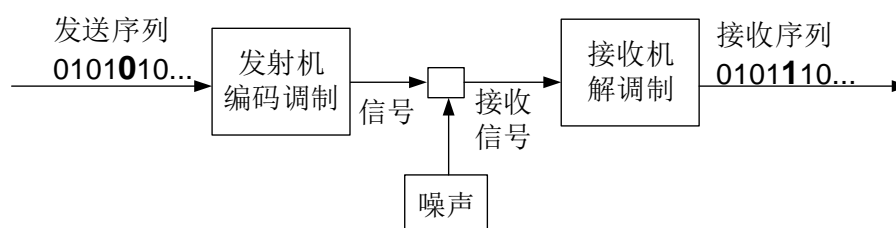


图 1-1 简化后的数字传输模型

二进制序列通常需要将 K 个比特作为一个“符号”进行传输，每个符号有 2^K 个不同状态。光传输利用光波的复振幅承载信号，因此可用复平面上不同的点来

对应不同的符号状态，这种将符号状态画在复平面上的图称为“星座图”，图上的点称为“星座点”。如图 2(a)所示的 QPSK（Quadrature Phase Shift Keying）调制，经过信道叠加噪声和接收机处理后，接收端的星座图不再是理想的四个点，而是会出现扩散。当接收机收到 1 个符号时，就将发送的符号判定为离该符号最近的星座点。显然，如果噪声过大，接收到的符号可能被判错从而产生误码，如图 2(b)中的蓝点。误码率（Bit Error Ratio, BER）定义为错误的比特数占总传输比特数的比例，例如传输了 50 个符号共 100 个比特，其中有 1 个符号被误判为相邻的符号，错误了 1 个 bit，则误码率为 0.01。BER 是衡量通信系统性能的最根本指标，采用纠错编码，只要纠前 BER 小于某个门限值（BER 容限值），纠错编码后就能实现纠后误码率为零的传输，本题中 BER 都是指纠错编码前的误码率（纠前 BER）。

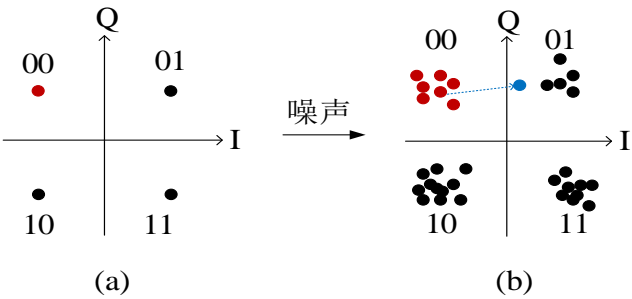


图 1-2 星座图与噪声导致误码的示意图

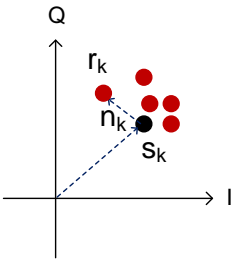


图 1-3 信号和噪声的相关定义示意图

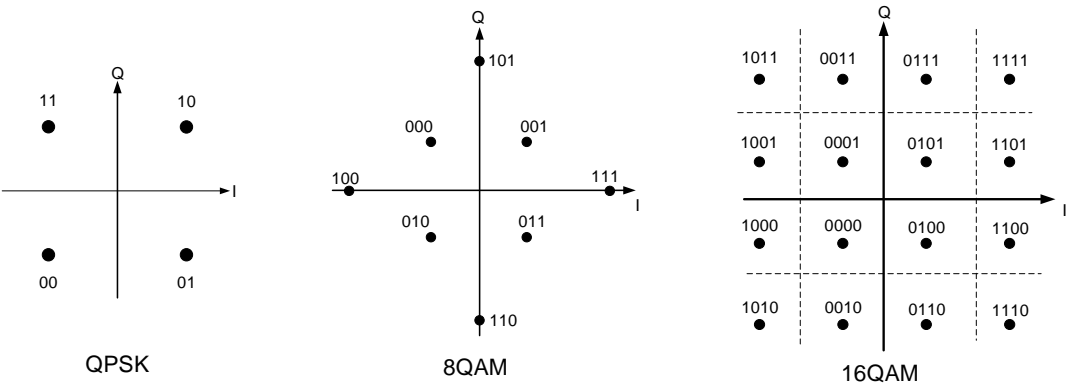


图 1-4 三种调制格式的编码方案

星座图的编码分布模式也称为调制格式，对于给定的调制格式，BER 和 SNR 呈一一对应的关系，纠前 BER 门限对应的 SNR 记做“SNR 容限点”。给出图 1-4 中所示的三种调制格式及编码方式（相邻星座点距离相等），每个符号等概率出现，分别称为 QPSK, 8QAM (Quadrature Amplitude Modulation, QAM), 16QAM。题目要求计算出 BER 与 SNR 的关系曲线，以及当 BER=0.02 时 SNR 容限点。

光传输链路由多个相同跨段的级联而成。如图 4 所示，几十 km 的光纤和一个放大器构成了 1 个跨段。信号每传输 15km，光功率衰减一半，经过一段光纤传输后，需要用放大器对光功率进行补偿。在信号、噪声同步放大的同时，放大器还引入自发辐射噪声；另一方面，光纤作为一种传输介质，其本身的非线性效应也会等效地引入噪声。其等效噪声功率与入纤功率近似呈平方关系，光纤功率为 1mW 时的非线性噪声约等于单个放大器噪声的 2/3。放大器的自发辐射噪声和光纤的非线性噪声都是加性噪声。

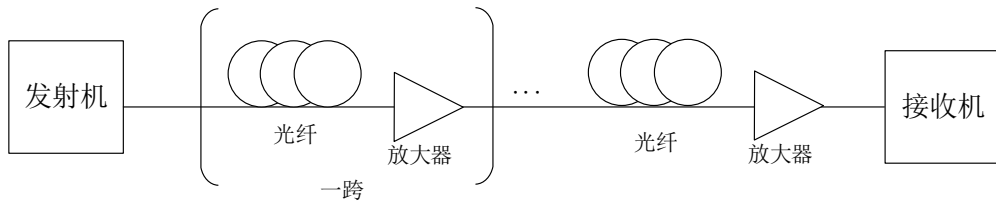


图 1-5 基本的光传输链路模型

本问题需要求解当单跨传输距离为 80km 和 100km 两种情况，以纠前误码率 0.02 为门限，图 1-5 给出的传输格式最远的传输距离(每跨距离×跨段数量)。

1.2.2 问题二：光传送网规划

表 1 给出进一步优化升级后的三种典型光传输设备参数。考虑到通信网络的目的是把更多的人更充分地连接到一起，我们按照如下方式定义网络的价值：

- 1) 每条直接连接两个城市/区域的链路当做 1 个连接，每个连接的价值定义为传输的容量与连接区域人口数的乘积（取两区域人口数乘积的 0.5 次方）
- 2) 网络的价值则是所有连接价值的加权和

$$\text{网络价值} = \sum \text{权重} * \text{容量} * \text{人口}$$

以图 5 给出的北京、南京、上海三座城市为例，若相互之间均互有连接，根据城市的距离可得出能传输的容量。若每条链路的权重为 1，进而再由人口算出网络价值（Network Value, NV）为

$$NV = \sqrt{21 \times 24} \times 16 \text{Tb/s} + \sqrt{21 \times 8} \times 16 \text{Tb/s} + \sqrt{24 \times 8} \times 32 \text{Tb/s} \approx 1010 \text{mTb/s}$$

其中 m 代表百万人(million)，Tb/s = 10^{12} bit/s，该网络的连接数为 3。

然而由于资源等因素制约，网络往往并不能让每对节点都直接连接，但可通

过中间节点连接起来。以图 5(b)为例，网络连接数为 2，北京和南京之间需通过上海中转。若北京上海之间的传输仅保留一半容量(100Gb/s)，而另一半容量用于南京到北京的信号传输(100Gb/s)，相应地南京与上海之间的直接传输容量也会降低至 300Gb/s，此时网络的价值为

$$NV = \sqrt{21 \times 24} \text{m} \times 8 \text{Tb/s} + \sqrt{21 \times 8} \text{m} \times 8 \text{Tb/s} + \sqrt{24 \times 8} \text{m} \times 24 \text{Tb/s} \approx 616 \text{mTb/s}$$

根据需要两个节点之间也可以有多个连接。

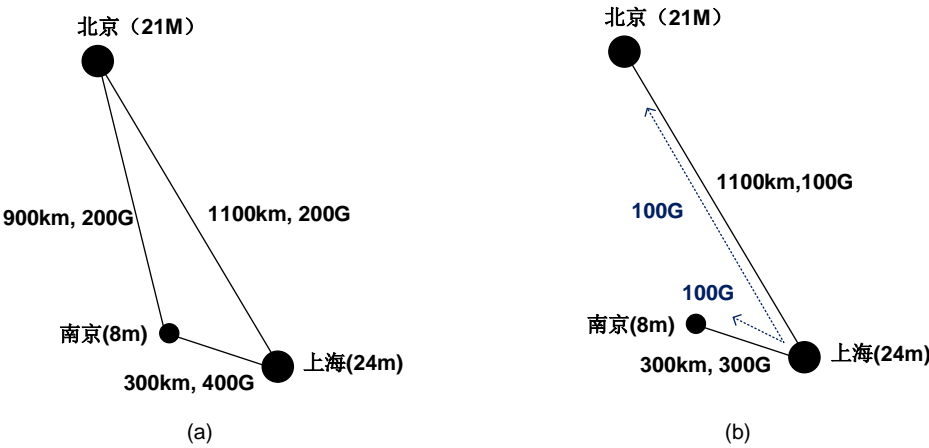


图 1-6 三个节点网络示意图

单波传输容量	最大传输距离	总容量
100 Gb/s	3000 km	8 Tb/s
200 Gb/s	1200 km	16 Tb/s
400 Gb/s	600 km	32 Tb/s

表格 1-1 不同传输格式的传输距离

请你们队考虑价值与需求为图 6 中的我国城市群制定光传送网规划，图中共有 12 个区域(其中北京/天津，深圳/广州均按 1 个区域对待)，

子问题-1:如果连接数从 16 增加到 33 条时，不考虑中间节点，给出你们的两个网络规划及其价值。网络价值最多是多少？

子问题-2: 存在中间节点，且两个节点之间可以有多个连接的情况下，重新解决子问题-1 并给出所有中间节点传输容量的分配，假定每条链路容量可任意分配，只要总容量不超过表 1 的规定。如果由市扩大为省（区）影响如何？（人口请从网上查找）

子问题-3:光传送网络价值有多个侧面，例如从运营商的角度，连接经济发达的地区会带来更多的收入，从政府的角度保障发展相对滞后地区的通信是均衡发展的要求等。你队认为制定光传送网络规划的目标函数应该是什么？前面制定的

规划有无变化？

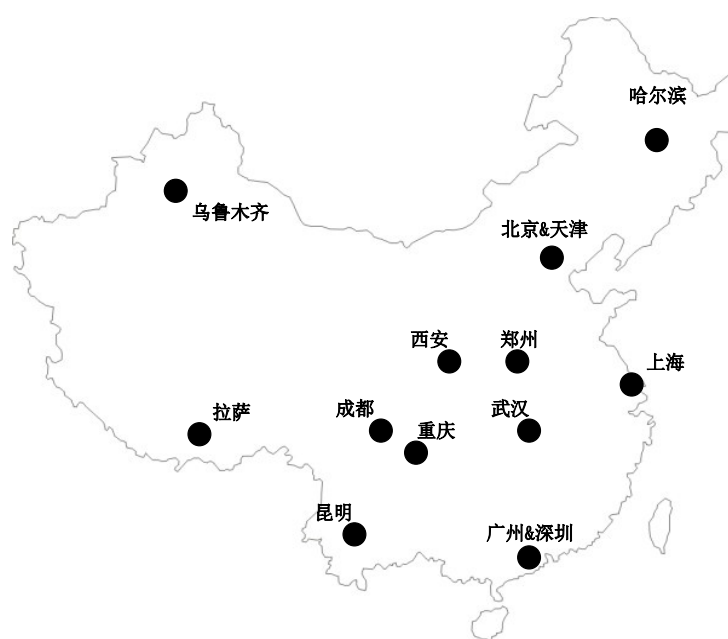


图 1-7 需要考虑的城市群

1.2.3 问题三：改善星座图

由第一问可知，纠前 BER 不变时，降低 SNR 容限点可以提高系统容忍噪声的能力，从而延长链路的总长度。本题要求尝试任意改变 16QAM 方案中星座点的位置、数量或每个点的概率，探索产生比图 1-4 中 8QAM（相邻各星座点之间距离相等）具有更低 SNR 容限点的调制方案？调制格式的信息熵需保持为 3bit。

信息熵定义为：

$$\Omega = -\sum_1^N p_k \log_2(p_k)$$

其中 p_k 为每个符号状态出现的概率，N 为状态数。图 5 所示的等概率情况下，QPSK、8QAM 和 16QAM 的信息熵分别为 2bit, 3bit 和 4bit。

2 问题一 光传送链路建模

2.1 模型假设

对于子问题（一）我们做出以下模型假设：

- (1) 对于不同的调制格式，信号的功率归一化为 1mW；
- (2) 噪声的分布在星座图上的分布符合二维高斯噪声模型；
- (3) 假设 QPSK、8QAM、16QAM 的星座图的理想星座点之间满足任意一符号与其最临近的星座点之间的距离相同的条件。

对于子问题（二）我们做出以下模型假设：

- (1) 光纤的入纤功率为 1mW；
- (2) 光纤在每跨中传输时的信噪比保持不变；

2.2 符号说明

符号	意义
P_s	信号的平均功率
P_n	噪声的平均功率
SNR	信号噪声功率比（信噪比）
N	调制方式的阶数，代表对应的星座图上的符号个数
K	星座图上的点 $K \in \{1, \dots, N\}$
\mathbf{s}_k	理想的接受到的符号向量
\mathbf{n}_k	噪声向量
\mathbf{r}_k	实际接收的符号向量
$PDF(k)$	星座点 k 的概率密度函数
D_k	星座点 k 的判决区域
d_{kj}	星座点 k 判决为星座点 j
$err(d_{kj})$	星座点 k 判决为星座点 j 时的错误 bit 数
P_{ASE}	放大器自发辐射噪声
P_{nl}	非线性效应引入的噪声功率
P_{s0}	初始信号功率
L	光传输链一跨的长度
Gain	每一跨中信号功率的增益系数
P_{N_add}	每一跨中增加的噪声信号
P_N^i	传输到第 i 跨时的总噪声功率

2.3 问题一(1) 纠前误码率与信噪比计算

2.3.1 分析与模型

该问题需要求解 BER 与 SNR 的关系曲线。在信号功率归一化（即 $P_s = \frac{1}{N} \sum_{k=1}^N |s_k|^2 = 1$ ）以及相邻最近符号间等距离的前提下，我们可以求出对应三种调制的符号坐标点。在一定的 SNR 水平下，我们可以根据公式

$$SNR = \frac{P_s}{P_n}$$

求出对应的噪声平均功率 P_n 。根据题目提供的条件，噪声服从均值为 0 的正态分布。该正态分布的方差 σ^2 等于噪声的平均功率。即：

$$\sigma^2 = P_n = \frac{1}{SNR} = 10^{\frac{-SNR(dB)}{10}}$$

在复平面上的二维正态分布的概率密度函数为：

$$PDF = f(\mathbf{x}) = \frac{1}{\sqrt{2\pi^2} |\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right]$$

其中 $\boldsymbol{\mu}$ 为均值， Σ 为协方差矩阵， ρ 为相关系数。因为在复平面 I、Q 方向上的噪声是独立的，所以 Σ 是对角矩阵，如下图所示，得到的二维高斯分布在 XOY 平面上的投影为圆。因此我们用均值为 s_k ，方差为 σ^2 的二维正态分布来表示星座图上 k 点的符号向量 r_k 的概率分布函数。

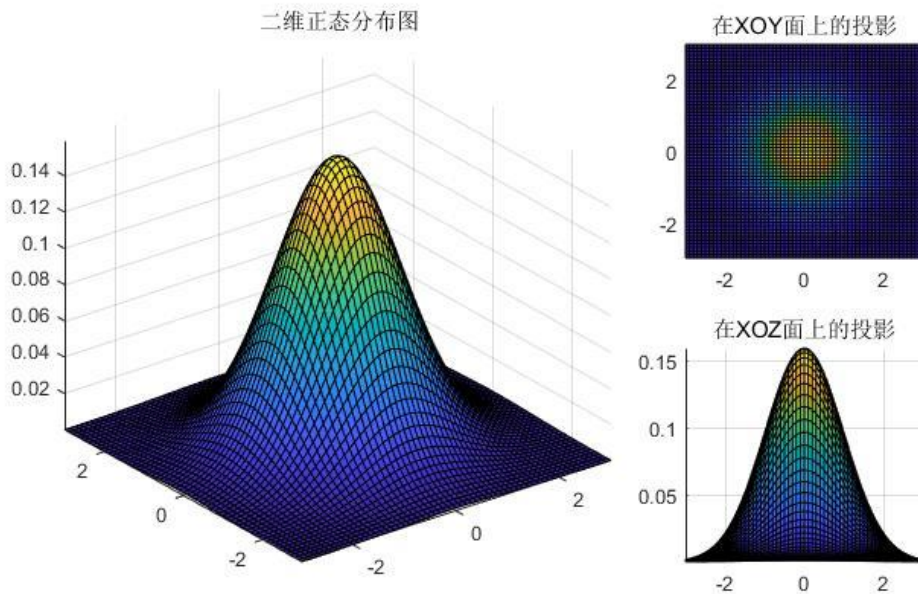


图 2-1 二维正态分布示意图

画出星座图中所有星座点之间的中垂线，中垂线分割出的区域为每一个星座点的判决区域。对于星座点 k 的概率分布函数， $\iint_{D_{k_j}} PDF(k_i) dx dy$ 表示符号 k_i 判决为 k_j 的概率，其错误 bit 数的平均期望是

$$\sum_{j=1}^N err(d_{k_j}) \times \iint_{D_{k_j}} PDF(k) dx dy$$

综合考虑所有的 k 点的概率分布，一个星座图的错误 bit 数期望为：

$$E_{bit} = \sum_{k=1}^N P(k) \times \sum_{j=1}^N err(d_{k_j}) \times \iint_{D_{k_j}} PDF(k) dx dy$$

这时可以计算 BER 为：

$$BER = \frac{E_{bit}}{\Omega}$$

2.3.2 问题求解

求解步骤分为以下几步：

Step1: 依据功率归一化公式求得具体的星座点坐标；

Step2: 依据星座点的位置划分判决区域；

Step3: 根据二维高斯分布的概率密度函数计算每一种判决 d_{ij} 的概率分布；

Step4: 结合星座点的概率 $P(k_i)$ 和概率密度函数得到错误比特数期望；

Step5: 计算设定 SNR 区间 BER；

Step6: 使用二分法不断逼近 $BER = 0.02$ 的 SNR 容限点；

2.3.2.1 QPSK

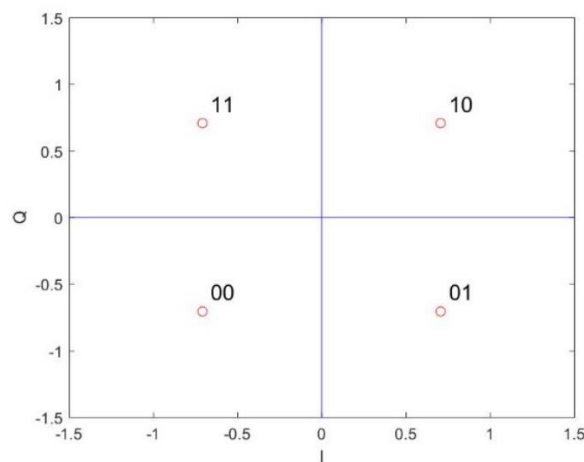


图 2-2 QPSK 调制星座图

如图 2.1 所示为 QPSK 调制的星座图, $N = 4$, 根据功率归一化得到星座点之间的距离 $d = \sqrt{2} \sqrt{mW}$ 。此时的高斯分布的方差可以计算为:

$$\sigma^2 = 10^{\frac{-SNR(dB)}{10}}$$

根据图中展示的编码方式, 每一个星座点的最近邻点有两个。因为不是格雷码编码, 其中一个是 2-bit 不同, 另一个是 1-bit 不同。根据四个星座点的的中垂线划分的判决区域有四个, 判决的错误 bit 数统计为表 2-1。可以看出四个星座点的判决错误类型是完全对称的, 即每一个星座点出错的期望是相同的。

	00	01	10	11
00	0	1	1	2
01	1	0	2	1
10	1	2	0	1
11	2	1	1	0

表格 2-1 QPSK 错误 bit 数

以 ‘00’ 为代表, 其概率密度分布函数为:

$$PDF_{qpsk} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\left(x + \frac{\sqrt{2}}{2}\right)^2 + \left(y - \frac{\sqrt{2}}{2}\right)^2}{2\sigma^2}\right)$$

计算 QPSK 的误码率 BER 为:

$$BER = \frac{1}{2} \left[\iint_{D_{0001}} PDF_{qpsk}(k_{00}) ds + \iint_{D_{0010}} PDF_{qpsk}(k_{00}) ds + 2 \right. \\ \left. * \iint_{D_{0011}} PDF_{qpsk}(k_{00}) ds \right]$$

在 Matlab 中根据公式计算出 BER 和 SNR 的曲线如图 2-2 所示。由于 BER-SNR 的函数是单调的, 通过二分法逼近得到较为精确的容限点 $SNR_{qpsk}(BER = 0.02) = 6.92$

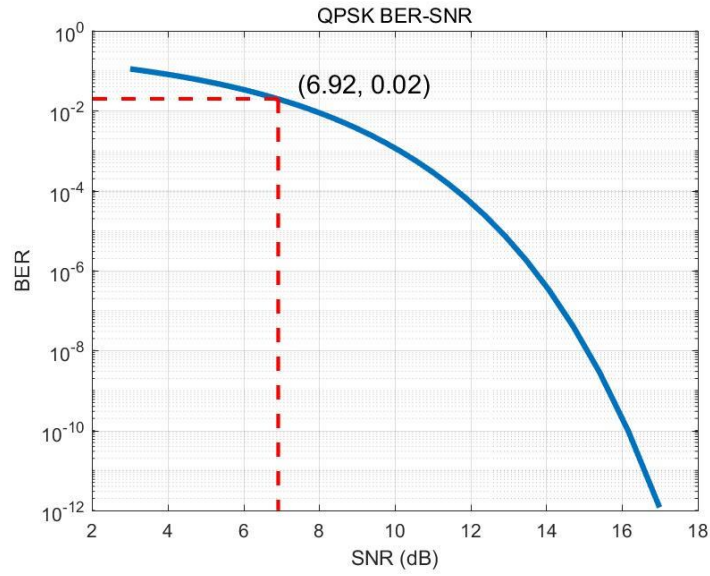


图 2-3 QPSK BER-SNR 关系曲线

2.3.2.2 8-QAM

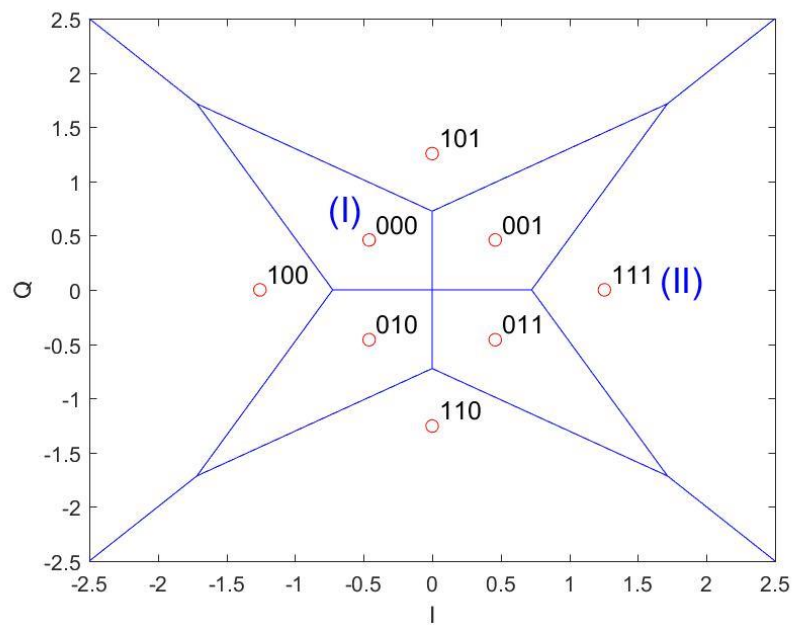


图 2-4 8QAM 调制星座图

如图 2.3 所示为 8QAM 调制的星座图， $N=8$ ，根据功率归一化得到星座点之间的距离 $d \approx 0.9194 \sqrt{mW}$ 。

根据图中展示的编码方式，可以将图中的点分为两类（I 和 II）。每个 I 类点

有四个距离为 d 的最邻近点，且其中必定有一个点是有两个 bit 不相同的；每个 II 类点仅有两个点为距离为 d 的最邻近点，且都符合格雷码的编码方式只有一个 bit 的不同。所以 I 类星座点的错误概率会大于 II 类点，因此需要分开计算其错误概率。而每一类的星座点的出错类型和出现概率都是对称的，因此可以用任意一点的计算结果来代表对应类型的 BER。

类型/星座点		I				II			
		000	001	011	010	111	101	110	100
I	000	0	1	2	1	3	2	2	1
	001	1	0	1	2	2	1	3	2
	011	2	1	0	1	1	2	2	3
	010	1	2	1	0	2	3	1	2
II	111	3	2	1	2	0	1	1	2
	101	2	1	2	3	1	0	2	1
	110	2	3	2	1	1	2	0	1
	100	1	2	3	2	2	1	1	0

表格 2-2 8QAM 错误 bit 数

对于 I 类点，以星座点（000）为例子，其概率密度函数为：

$$PDF_{8QAM_I} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\left(x + \frac{d}{2}\right)^2 + \left(y - \frac{d}{2}\right)^2}{2\sigma^2}\right)$$

其错误 bit 数的期望为：

$$\begin{aligned} E_{bit}^I &= 2 \times \iint_{D_{100}} PDF_{8QAM_I} ds + 2 \times \iint_{D_{010}} PDF_{8QAM_I} ds \\ &\quad + 2 \times \iint_{D_{110}} PDF_{8QAM_I} ds + \iint_{D_{011}} PDF_{8QAM_I} ds \end{aligned}$$

对于 II 类点，以星座点（111）为例子，其概率密度函数为：

$$PDF_{8QAM_II} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\left(x + \frac{d}{2}\right)^2 + \left(y - \frac{d}{2}\right)^2}{2\sigma^2}\right)$$

其错误 bit 数的期望为：

$$\begin{aligned} E_{bit}^{II} &= 2 \times \iint_{D_{001}} PDF_{8QAM_II} ds + 2 \times \iint_{D_{101}} PDF_{8QAM_II} ds \\ &\quad + 2 \times \iint_{D_{000}} PDF_{8QAM_II} ds + \iint_{D_{100}} PDF_{8QAM_II} ds \end{aligned}$$

最后 BER 的计算公式为：

$$BER_{8QAM} = \frac{\frac{8}{16} \times E_{bit}^I + \frac{8}{16} \times E_{bit}^{II}}{\log_2 N} = \frac{1}{6} \times E_{bit}^I + \frac{1}{6} \times E_{bit}^{II}$$

在 Matlab 中根据公式计算出 BER 和 SNR 的曲线如图 2-4 所示。由于 BER-SNR 的函数是单调的，通过二分法逼近得到较为精确的容限点 $\text{SNR}_{8\text{QAM}}(\text{BER} = 0.02) = 10.341$

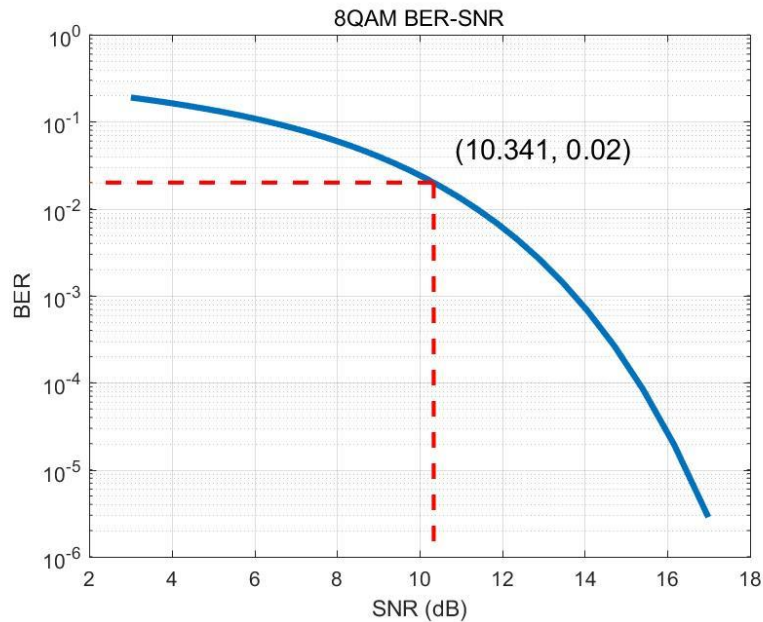


图 2-5 8-QAM BER-SNR 关系曲线

2.3.2.3 16-QAM

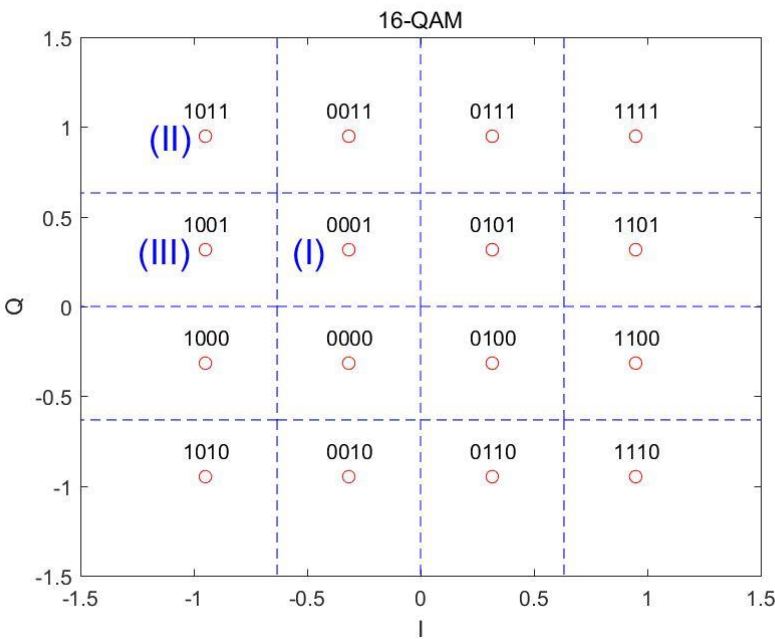


图 2-6 16-QAM 调制星座图

如图 2.3 所示为 16-QAM 调制的星座图， $N = 16$ ，根据功率归一化得到星座点之间的距离 $d \approx \sqrt{\frac{2}{5}} \sqrt{mW}$ 。此时的高斯分布的方差可以计算为：

$$\sigma^2 = 10^{\frac{-SNR(dB)}{10}}$$

根据图中展示的编码方式，可以将图中的点分为三类（I~III）。上图中的调制模式采用格雷码编码，因此每一类和最邻近点的差别只有 1bit，但是每类点的最邻近点的个数是不相同的。因此需要同 8QAM 一样分类计算每一类点的错误比特数的期望，最后结合每一类星座点出现的概率综合计算 BER。采用和上两个小节相同的求解方法。在 Matlab 中根据公式计算出 BER 和 SNR 的曲线如图 2-6 所示。由于 BER-SNR 的函数是单调的，通过二分法逼近得到较为精确的容限点 $SNR_{16QAM}(BER = 0.02) = 12.7726$

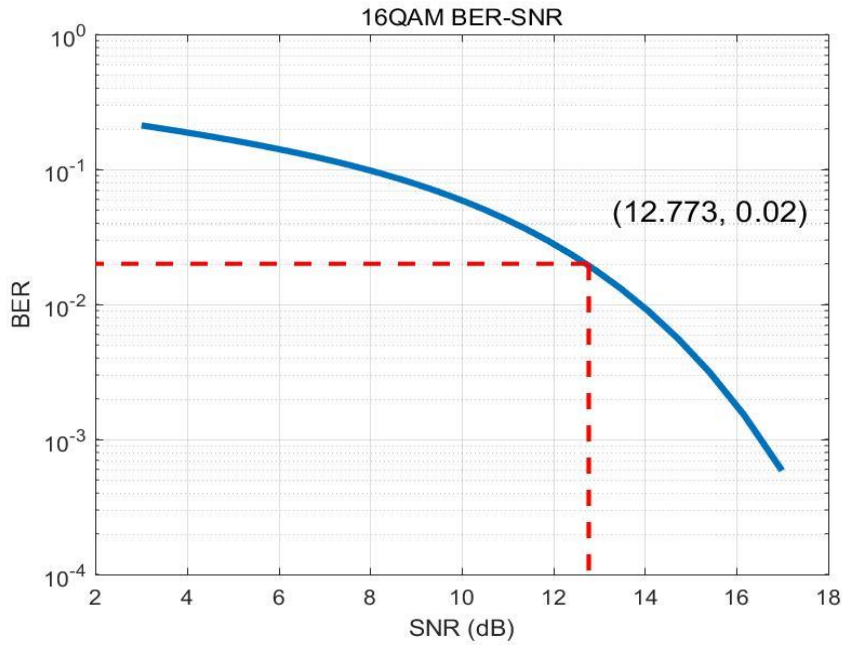


图 2-7 16-QAM BER-SNR 关系曲线

2.3.3 结果

三种调制格式的编码方案的结果汇总如下图所示

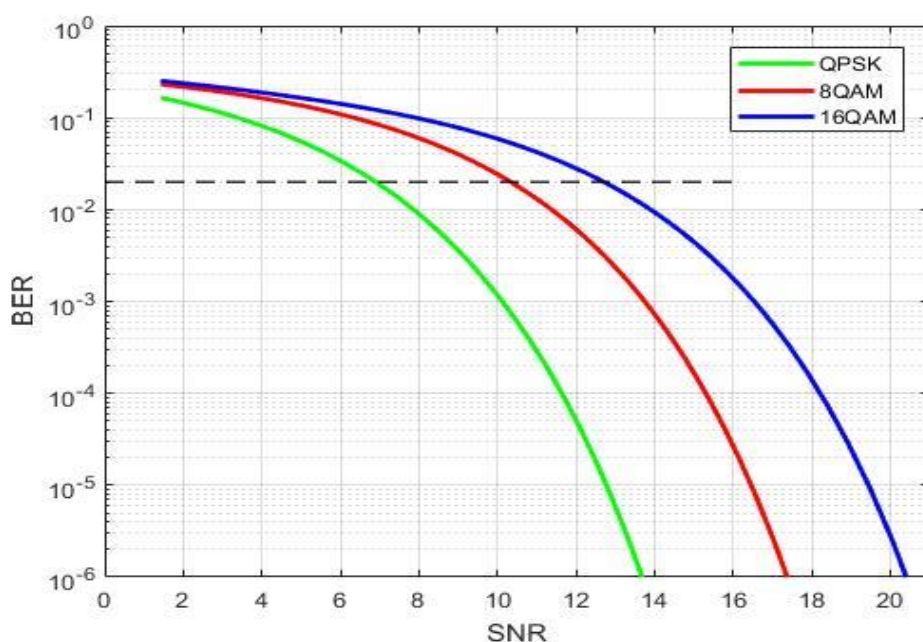


图 2-8 三种调制方式的 BER-SNR 性能曲线

其中三种调制的 BER = 0.02 的容限点分别为：

调制	QPSK	8QAM	16QAM
SNR 容限点	6.9207	10.341	12.7726

2.4 问题一(2) 光链路性能计算

2.4.1 问题分析

该题需模拟光传输链路的信号噪声模型。光传输链路由多个相同跨段的级联而成。如图 4 所示，几十 km 的光纤和一个放大器构成了 1 个跨段。信号每传输 15km，光功率衰减一半，经过一段光纤传输后，需要用放大器对光功率进行补偿。

一方面，在信号、噪声同步放大的同时，放大器还引入自发辐射噪声 (Amplified Spontaneous Emission Noise 简称 ASE)，公式为

$$P_{ASE} = 2\pi h f B (NF - 1/Gain)$$

另一方面，光纤作为一种传输介质，其本身的非线性效应也会等效地引入噪声。其等效噪声功率与入纤功率近似呈平方关系，光纤功率为 1mW 时的非线性噪声约等于单个放大器噪声的 2/3。放大器的自发辐射噪声和光纤的非线性噪声都是加性噪声。非线性噪声在每跨开头叠加，放大器噪声在每跨末尾叠加；每跨中间信噪比保持不变。每跨叠加的噪声间相互独立。

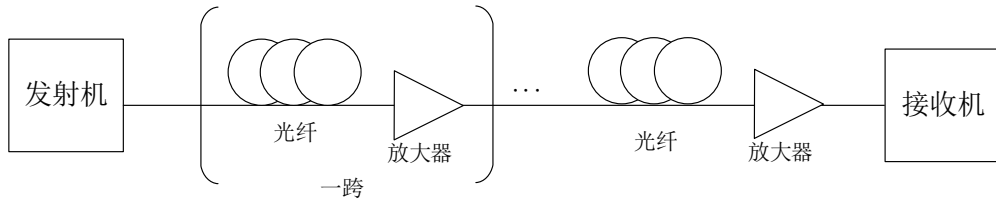


图 2-9 基本的光传输链路模型

首先分析光信号功率的变化。信号传输 15km，光功率就衰减一半，这是一种典型的指数函数变化规律，因此可以将传播光信号的功率表示为：

$$P_s = P_{s0} \times 2^{\frac{-L}{15}}$$

当跨长 L 固定为 80km 或 100km 时， P_s 的衰减相对固定，为 $\text{Loss} = 2^{\frac{-L}{15}}$ ，相应的信号功率的增益为：

$$\text{Gain} = 2^{\frac{L}{15}}$$

然后分析噪声功率的变化，根据题意信号的噪声分为两种，一种是自发辐射噪声 P_{ASE} ，另一类是非线性噪声 P_{nl} ，其中

$$P_{ASE} = 2\pi h f B \left(NF - \frac{1}{\text{Gain}} \right)$$

根据文献，光纤信号的常见功率约为 1mW，且从信号开始传输到传输结束，光信号功率一直占据光链路的主要部分，因此根据题目，以下关系始终成立

$$P_{nl} = \frac{2}{3} P_{ASE}$$

由于功率噪声是加性噪声，所以总噪声功率为

$$P_{nl} + P_{ASE} = \frac{5}{3} P_{ASE}$$

由于噪声与信号在信道中同步衰减并在放大器中得到同步的增益，因此在信号功率恢复到 P_{s0} 的同时，噪声功率得到了 Gain 倍的增益，即

$$P_{N_add} = \frac{5}{3} \text{Gain} \times P_{ASE}。$$

在不同跨间的信号传输中，信号的功率保持不变为 P_{s0} 。设 P_N^i 为信号经过第 $i, i \in \{1, 2, \dots\}$ 跨的链路传输之后的噪声，而信道中传输的噪声 P_N 不断地叠加，前一跨的噪声经过一跨的衰减之后又受到 Gain 倍功率增益。因此上一跨传输的噪声功率 P_N^{i-1} 经过信道和放大之后保持不变，第 i 跨输出的噪声只是上一跨的噪声加上 P_{N_add} ，即

$$P_N^i = P_N^{i-1} + P_{N_add}$$

当 $\text{SNR} = 10 \log_{10} \left(\frac{P_{s0}}{P_n^i} \right) < \text{SNR}_{\text{limit}}$ 光链路中的信号不再能保证纠前误码率 $\text{BER} > 0.02$ ，此时，不再能保证光信号传输的稳定。

2.4.2 模型与求解

本题的模型如下：

约束条件：

$$P_{s0} = 1; \quad P_N^0 = 0; \quad \text{Gain} = 2^{\frac{l}{15}}$$

$$P_{ASE} = 2\pi h f B \left(NF - \frac{1}{\text{Gain}} \right)$$

$$P_{Nadd} = \frac{5}{3} \text{Gain} \times P_{ASE}$$

$$P_N^i = P_N^{i-1} + P_{Nadd}$$

求解

$$\text{求最小的 } i \in \{0, 1, 2, \dots\} \text{ 使得 } \text{SNR} = 10 \log_{10} \left(\frac{P_{s0}}{P_n^i} \right) < \text{SNR}_{\text{limit}}$$

该问题的求解比较直接：

- 1、使用 Matlab 工具，对上述的问题建模。
- 2、初始化信号噪声功率的初值，迭代计算出每一跨结束时候的功率信噪比 SNR，
- 3、如果信噪比低于问题一(1)中的信噪比容限，则停止迭代。
- 4、根据跨长和跨长度求出最远传输距离。

2.4.3 结果与分析

	跨长80km			跨长100km		
	QPSK	8QAM	16QAM	QPSK	8QAM	16QAM
最多跨数	17	8	4	6	3	1
最远传输距离	1360	640	320	600	300	100

表格 2-3 最远传输距离求解结果

由表 2-3 的数据可以看出跨长越短，最远传输距离和最多跨数更多。这符合实际情况，因为在跨中信号的衰减时指数形式的，较短的跨长使得跨两端数据的增益不至于过大。引入的噪声水平就相对较小，因此有利于信号的长距离传输。

同时对于不同的调制方式，调制阶数越小，最远传输距离越远。这是由于相同 BER 容限，较小的调制阶数的 SNR 容限更低，自然带来更远的传输距离。

3 问题二 光传网规划

3.1 模型假设

考虑到通信网络的目的是把更多的人更充分地连接到一起,我们按照如下方式建立网络的模型:

- 3) 每条直接连接两个城市/区域的链路当做 1 个连接, 每条连接都默认是同一水平面上的直线相连, 忽略地形等环境因素的影响。
- 4) 每个连接的价值定义为传输的容量与连接区域人口数的乘积 (取两区域人口数乘积的 0.5 次方)
- 5) 网络的价值则是所有连接价值的加权和

$$\text{网络价值} = \sum \text{权重} * \text{容量} * \text{人口}$$

4) 在给定了城市人口, 相互距离之后, 城市间光传送网问题就简化成带权重的连通图问题。设 $G = (V, E, W)$, 其中 V 是 12 个区域顶点集, E 是所有连接边集, W 是所有对应边的权重即连接价值集。我们需要做的就是求得一组边的集合 $Z (Z \subset E)$, 使得连通图的总价值最大, 也即

$$\max \sum w_e, w \in W, e \in E$$

3.2 符号说明

符号	意义
G	带权重的连通图
V	12 个区域顶点集
W	边的权重即连接价值集
E	所有连接边集
Z	E 的子集
w	W 中的元素
e	E 中的元素
P	存放最大生成树的顶点集合
Q	存放最大生成树的边集合
A, B, C	三个不同城市的人口数量
l, m, n	三个城市间的传输容量
$w_{i,j}$	i, j 之间通信总流量
$p(i), p(j)$	i, j 两城市的人口数
$m(i, j), n(i, j)$	将某一棵树中的 i, j 边去掉后所产生的两棵子树
$\min bw$	两城市间通信的最小带宽
E_i	i 节点单独作用产生的损失
$E_{i,j}$	将连两个点 i, j 联合作用所产生的损失

C	i,j之间相连的信道容量
μ	i,j之间相连的地形评分
$D_{i,j}$	i,j城市的距离
$P_{i,j}$	i,j城市的人口
$G_{i,j}$	i,j城市的 GDP
D_{max}	可接受最长距离（这里设为 3000）
a, δ	皆可调参数（这里分别设为 100, 10^6 ）
α, β, γ	皆为可调参数（这里分别设为 0.4, 0.5, 2）

3.3 问题二(1) 无中间节点规划

3.3.1 问题分析

问题给出 12 个区域（如图所示）以及不同距离的传输容量（如表所示），不考虑中间节点且两连接点间只允许一条连接，分别求 16 条和 33 条连接的最大网络价值。

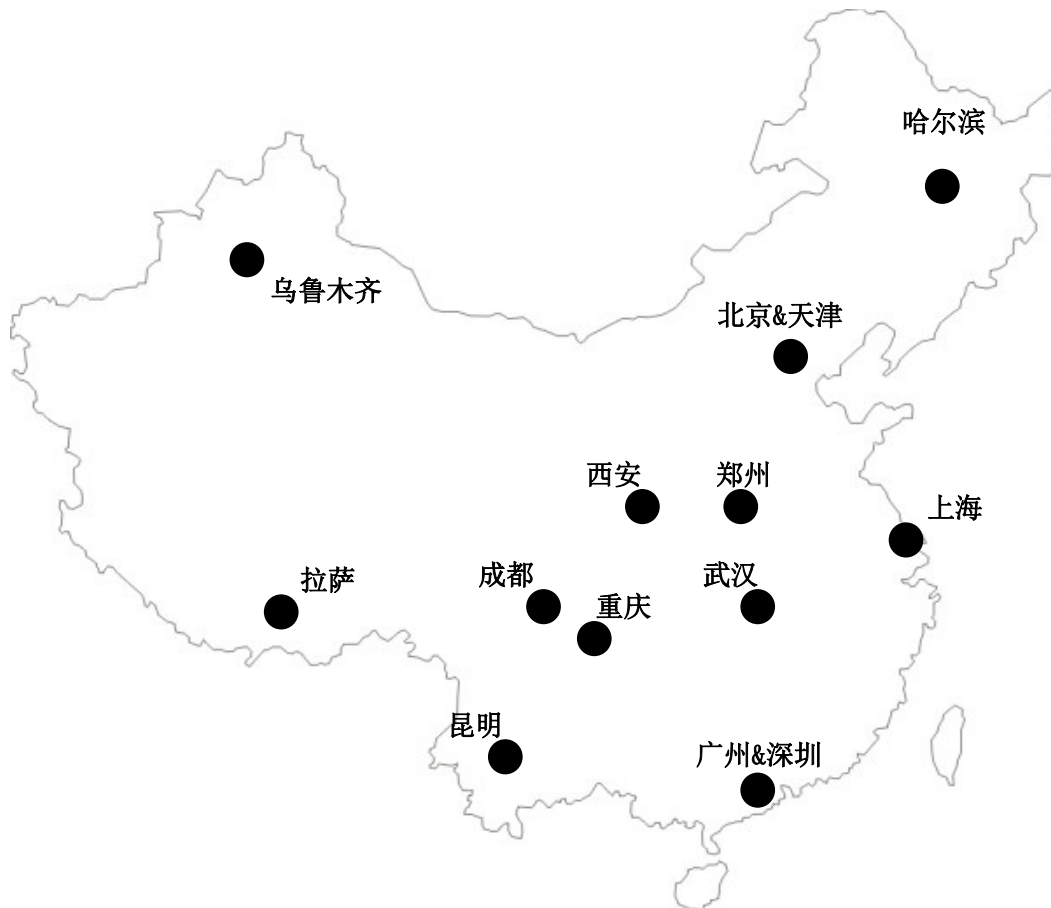


图 3-1 给定城市群

单波传输容量	最大传输距离	总容量
100 Gb/s	3000 km	8 Tb/s
200 Gb/s	1200 km	16 Tb/s
400 Gb/s	600 km	32 Tb/s

表格 3-1 不同距离的传输容量

任意一条连接的价值 = 权重 * 容量 * 两区域人口的几何均值（本问权重为 1），在不考虑中间节点且两连接点间只允许一条连接的情况下，城市之间两两相连的价值都可以求出。题目要求所有城市必须连入城市连通网，所以本题就是在城市两两连接之间找出最大生成树的问题。再在最大生成树(12 个区域只有 11 条连接)的基础上，添加剩下的最大价值连接，使得满足问题的 16 条和 33 条连接的要求。

3.3.2 问题求解

从网上查得 12 个区域人口以及两两之间的距离，计算出城市的连接价值，如下表所示。

	户籍人口(万人)
哈尔滨	955
北京&天津	3727
上海	2418
郑州	988
武汉	1089
西安	845
重庆	3075
成都	1604
昆明	480
广州&深圳	2703
拉萨	90
乌鲁木齐	267

表格 3-2 区域人口(2016 人口普查)

	哈尔滨	北京&天津	上海	郑州	武汉	西安	重庆	成都	昆明	广州&深圳	拉萨	乌鲁木齐
哈尔滨	0	1038	1677	1653	2000	1948	2506	2574	3139	2788	3556	3040
北京&天津	1038	0	1064	620	1052	900	1450	1526	2103	1880	2567	2412
上海	1677	1064	0	830	685	1230	1441	1657	1960	1200	2900	3265
郑州	1653	620	830	0	466	440	883	1383	1500	1292	2183	2443
武汉	2000	1052	685	466	0	654	758	987	1293	835	2230	2767
西安	1948	900	1230	440	654	0	578	620	1200	1320	1760	2118
重庆	2506	1450	1441	883	758	578	0	261	652	976	1500	2316
成都	2574	1526	1657	1383	987	620	261	0	637	1230	1250	2065
昆明	3139	2103	1960	1500	1293	1200	652	637	0	1100	1287	2530
广州&深圳	2788	1880	1200	1292	835	1320	976	1230	1100	0	2540	3300
拉萨	3556	2567	2900	2183	2230	1760	1500	1250	1287	2540	0	1622
乌鲁木齐	3040	2412	3265	2443	2767	2118	2316	2065	2530	3300	1622	0

表格 3-3 城市间距离(单位 km)

	哈尔滨	北京&天津	上海	郑州	武汉	西安	重庆	成都	昆明	广州&深圳	拉萨	乌鲁木齐
哈尔滨	0	30185.71	12156.82	7770.879	8158.412	7186.543	13709.27	9901.337	0	12853.3	0	0
北京&天津	30185.71	0	48031.69	30702.81	32233.96	28394.1	27082.72	19560.14	10700.16	25391.75	4633.306	7980.412
上海	12156.82	48031.69	0	24730.14	25963.43	11435.27	21814.27	15755.07	8618.64	20452.25	3731.981	0
郑州	7770.879	30702.81	24730.14	0	33192.69	29238.62	27888.23	10070.95	5509.207	13073.49	2385.557	4108.886
武汉	8158.412	32233.96	25963.43	33192.69	0	15348.37	29279.02	21146.38	5783.95	27450.92	2504.524	4313.796
西安	7186.543	28394.1	11435.27	29238.62	15348.37	0	51582.32	18627.33	5094.939	12090.42	2206.173	3799.916
重庆	13709.27	27082.72	21814.27	27888.23	29279.02	51582.32	0	71068.1	19438.52	46128.1	4208.563	7248.834
成都	9901.337	19560.14	15755.07	10070.95	21146.38	18627.33	71068.1	0	14039.21	16657.71	3039.579	5235.375
昆明	0	10700.16	8618.64	5509.207	5783.95	5094.939	19438.52	14039.21	0	18224.84	1662.769	2863.955
广州&深圳	12853.3	25391.75	20452.25	13073.49	27450.92	12090.42	46128.1	16657.71	18224.84	0	3945.793	0
拉萨	0	4633.306	3731.981	2385.557	2504.524	2206.173	4208.563	3039.579	1662.769	3945.793	0	1240.129
乌鲁木齐	0	7980.412	0	4108.886	4313.796	3799.916	7248.834	5235.375	2863.955	0	1240.129	0

表格 3-4 城市连接价值(两城市距离超过 3000km 置零)

依照上述分析，本问题先按照求解最大生成树的问题进行求解。连通赋权图 $G = (V, E, W)$ ，其中 V 是 12 个区域顶点集， E 是所有连接边集， W 是所有对应边的权重即连接价值集。设集合 P 和 Q ， P 用来存放最大生成树的顶点， Q 用来存放最大生成树的边。

Prim 算法如下：

- (1) $P = \{v_i\}$, $Q = \emptyset$, 其中 $v_i \in V$;
- (2) While $P \neq V$

找最大边 pv , 其中 $p \in P$, $v \in V - P$;

$P = P + \{v\}$;

$Q = Q + \{pv\}$;

End

因为任意两个城市之间互联的价值都不相同，所以，Prim 算法产生的最大生成树唯一且为全局最优解。根据算法求得的最大生成树如下图所示。

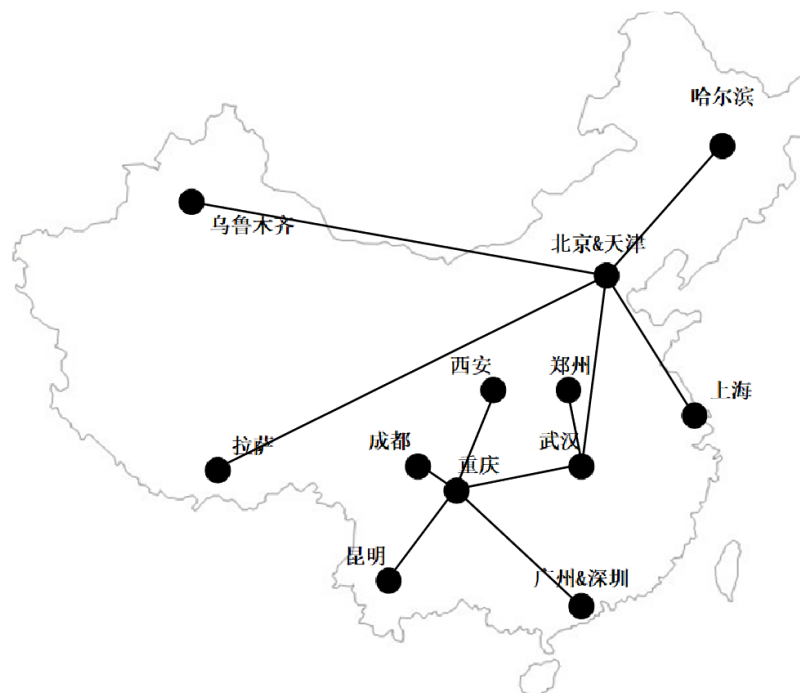


图 3-2 最大生成树连接图

在最大生成树基础上，在剩下的连接边集合里选出剩余的拥有最大连接价值的连接，可以分别得到 16 条连接和 33 条连接的最大连接价值网络。

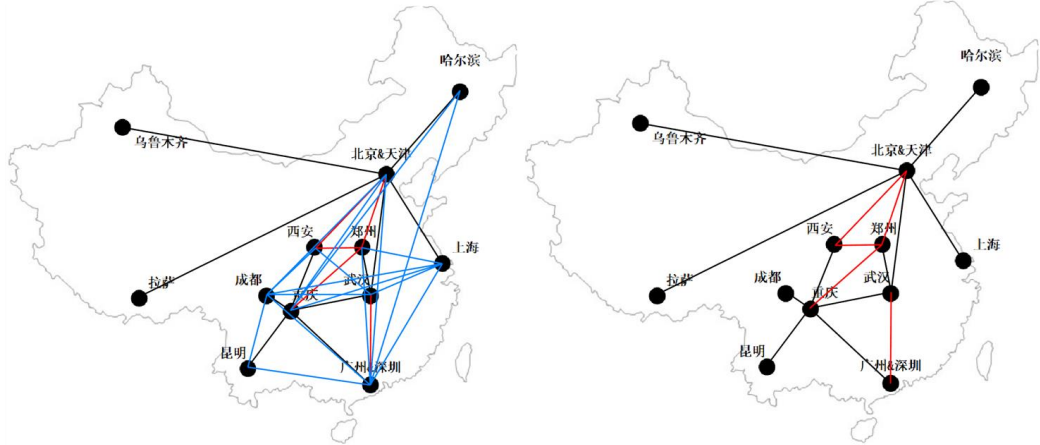


图 3-3 16 条连接和 33 条连接的最大连接价值网络

最大网络价值分别为 5174.3mTb/s 和 8418.6mTb/s。

3.4 问题二(2) 有中间节点规划

3.4.1 问题分析

问题二(2)比问题二(1)多加入了中间节点的分流机制，在这种分流机制的影响下，有可能会网络结构相同但通过改变不同的链路容量达到比不利用分流机制获得更高网络价值的情况。所以要对整体网络加以整体优化得到最优解。

但是对于整个问题来说，由于不同的网络结构会有不同的规划模板，在不同的模板下又会有更多的优化方案，所以在整体优化时把所有要素都考虑进去是很复杂的。

因为问题二(1)生成的最大生成树是全局最优解（没有重复权值的边），同时引入中间节点机制后也是求网络最大价值，所以接下来的节点分流最优解很可能在最大生成树的基础上产生。对此，我们需要进一步验证。

接下来就是考虑如何连接 11 条连接之后其他的连链接。我们需要先考察图上有哪些结构是引入分流机制比没有引入分流机制的价值更大，再将此结构与最大生成树结构相比，是否将连接线加在这个结构上比将连接线加在最大连接价值边上收益更大。如果最大连接价值边的收益更大，那么接下来的所有连接线将重复加在这个价值最大边上（本题允许城市间可以有多个链接）。相反，则重复加在分流结构上。

3.4.2 问题求解

按照以上分析的思路，我们本题求解分成两步。首先，我们分析最大生成树中的 11 个连接如何考虑有中间节点机制后分配流量使整体网络价值达到最大，之后我们分析在有了树形结构之后的其余连接如何分配能使总体价值最大化。

- 1) 证明城市互连结构里，即使考虑分流机制，也是最大生成树价值最大。
考虑如下图所示三城市结构。

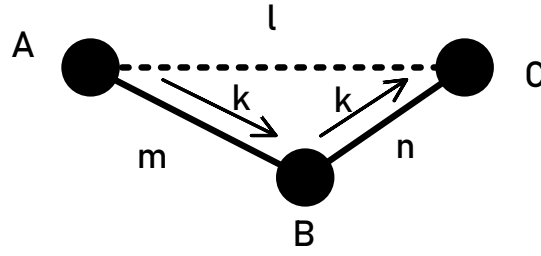


图 3-4 三城市连接结构

A, B, C 表示三个不同城市的人口数量，l, m, n 表示三个城市间的传输容量，那么不考虑中间节点机制连接的网络价值总量为

$$m\sqrt{AB} + n\sqrt{BC} \quad (3.1)$$

假设 AB 间有 k 的容量以分流流量的形式分配给 A 到 C 的连接，那么考虑中间节点机制连接之后的网络价值总量为

$$k\sqrt{AC} + (m - k)\sqrt{AB} + (n - k)\sqrt{BC} \quad (3.2)$$

将 (3.2) 式减去 (3.1) 式并化简，得中间节点机制的引入导致的网络价值总量提升为

$$k(\sqrt{AC} - \sqrt{AB} - \sqrt{BC}) \quad (3.3)$$

如果 (3.3) 式大于零，说明引入分流机制比直接互联价值更大。定性分析 (3.3) 式我们可以看出，在 A, C 较大而 B 较小时此式更有可能取得正值。而分析实际数据可以看出，人数比较小的城市（< 500 万人）只有拉萨、乌鲁木齐和昆明，和这三座城市相连的所有边收益都不是很大，所以如果用这三座城市作为中间节点，网络则会相比最大生成树损失较多价值。所以 (3.3) 式所引入的额外价值很有可能并没有太多实际意义。

下面我们进行定量分析，采用 matlab 编程的方式，遍历了所有可能的三城市互联结构，发现确实只有在以乌鲁木齐、拉萨或者昆明作为中间节点时，引入分流机制才能带来价值提升，而且最大的价值提升仅有 26690（分流额外价值+路线本身价值），小于最大生成树里除这三个边缘城市的连接外最低价值边 29279。也就是说我们用最大生成树中除连接乌鲁木齐、拉萨或者昆明的边中权重最小的边来产生能使分流机制带来最大收益的三元结构时都不能使总价值比最大生成树的总价值大。

三城市互联也仅有三个边缘城市能有提升，那么四或者四以上城市互联引入分流机制能获取更大价值的概率更小。

为了进一步验证，我们在最大生成树的基础上进行了建模验证。算法如下：

$$\text{目标函数} \quad \text{Max:} \quad E = \sum_{i,j} w_{i,j} \sqrt{p(i)p(j)} \quad (3.4)$$

其中 $w_{i,j}$ 代表 i, j 之间通信总流量， $p(i)$ 和 $p(j)$ 为 i, j 两城市的人口数。

$$\text{约束} \quad \sum_{m(i,j),n(i,j)} w_{m(i,j),n(i,j)} = bw_{i,j} \quad (3.5)$$

$$V_{i,j} w_{i,j} > \min bw \quad (3.6)$$

其中 $m(i,j)$ 和 $n(i,j)$ 可以看成是将某一棵树中的 i,j 边去掉后所产生的两棵子树。 $bw_{i,j}$ 为 i,j 两节点之间总带宽。该式相当于累加边 i,j 两边的任意一对节点的通信流量。 $\min bw$ 是两城市间通信的最小带宽，以保证城市间两两都有专用通信容量。

令 $\min bw = 0$ ，最后解得就是最大生成树。所以，最大生成树也是引入分流机制的树中，网络连接价值最大的。

2) 证明添加成树之外的边时三元结构中引入分流机制所得价值比当前最大价值要低。

根据 1) 里所说，只有乌鲁木齐、拉萨和昆明三个城市作为中间节点，引入分流机制才会有价值提升。考虑第 12 条连接用于构成以这三个城市为中心的三元结构，那么所有的可能里最大的价值提升为 26690，比重庆和成都之间的价值 71068 更低，所以为价值最大化考虑，第 12 条连接重复加在重庆和成都之间能获得最大的连接价值。

3) 证明添加成树之外的边时任意元结构中引入分流机制所得价值比当前最大价值要低。

在考虑任意个节点相连并引入分流机制时，我们计算了所有可能的连接结构中，能使得分流机制引入额外价值可以与成都重庆之间连接的价值等价所需要的分流带宽。在三元结构时所需最小带宽为 31.45，四元结构为 40.65，随着中间节点的增多，所需要的带宽在不断增大。而经过验证发现所有的这种可以通过分流机制得到提升的组合中，中间节点中一定会出现乌鲁木齐、拉萨和昆明三个城市中的至少一个，所以添加一条线路所能带来的该结构整个链路最大也只能是 $8 + 8 = 16$ 。

综合考虑 1)、2)，城市两两之间必须要有通信容量，这里取 $\min bw = 0.1\text{TB/s}$ ，算得前 11 条边的通信容量分配如下表所示，后 5 条和后 22 条重复加在重庆和成都之间，就能获得最大的网络价值。

	哈尔滨	北京&天津	上海	郑州	武汉	西安	重庆	成都	昆明	广州&深圳	拉萨	乌鲁木齐
哈尔滨	0	15	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
北京&天津		0	15	0.1	12.6	0.1	0.1	0.1	0.1	0.1	7	7
上海			0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
郑州				0	31	0.1	0.1	0.1	0.1	0.1	0.1	0.1
武汉					0	0.1	12.6	0.1	0.1	0.1	0.1	0.1
西安						0	31	0.1	0.1	0.1	0.1	0.1
重庆							0	31	15	15	0.1	0.1
成都								0	0.1	0.1	0.1	0.1
昆明									0	0.1	0.1	0.1
广州&深圳										0	0.1	0.1
拉萨											0	0.1
乌鲁木齐												0

表格 3-5 城市间互相通信的分配容量

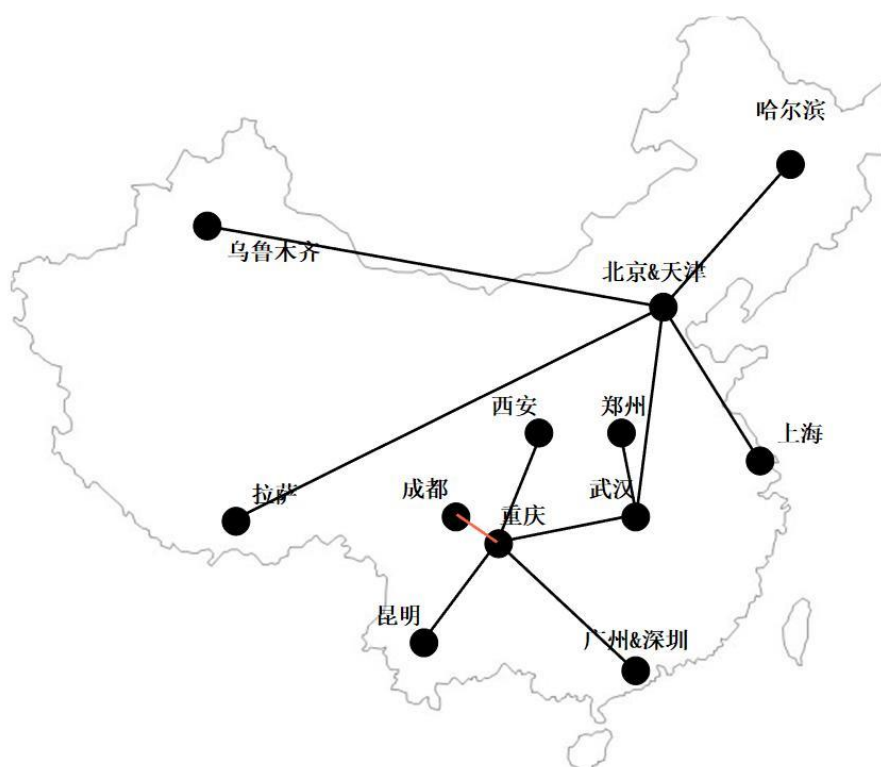


图 3-5 16 条连接和 33 条城市连接图，红色分别表示重复连接 5 次和 17
16 条连接和 33 条连接的总价值分别为 7069.9mTb/s和 19151.46mTb/s。
考虑省人口，如下表：

	人口(万)
哈尔滨	3799
北京&天津	3735
上海	2420
郑州	9532
武汉	5885
西安	3813
重庆	3048
成都	8262
昆明	4771
广州&深圳	10999
拉萨	331
乌鲁木齐	2398

表格 3-6 省人口(2016 人口普查)

将省人口代入，然后重做本题，得到 16 条连接和 33 条连接的总价值分别为 23510.5mTb/s和 64254.4mTb/s，连接网络如下

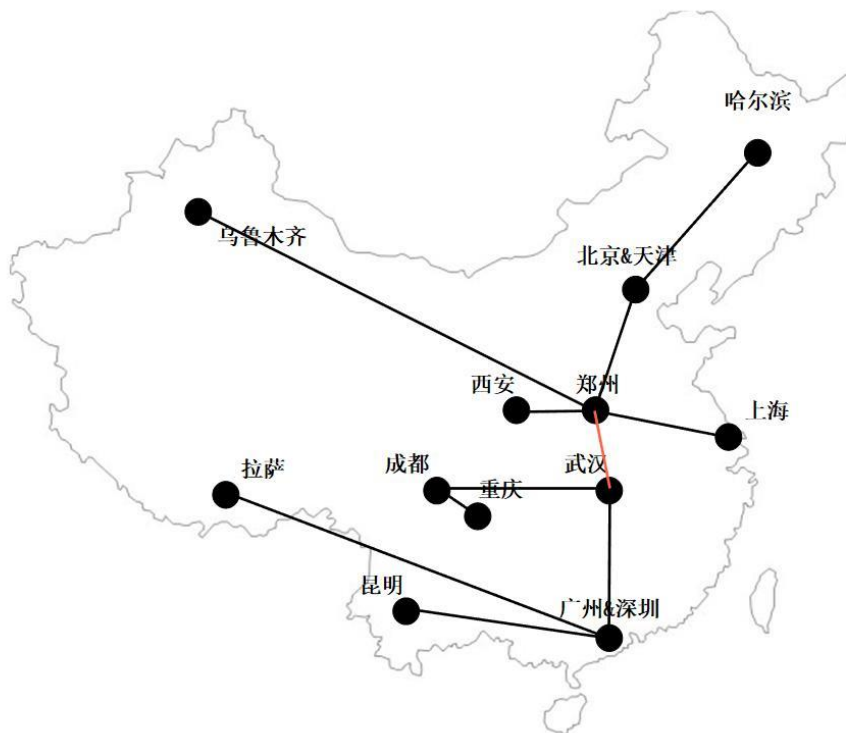


图 3-6 16 条连接和 33 条省连接图，粉色分别表示重复连接 5 次和 17 次

3.5 问题二(3) 重新设计目标函数

现有的目标函数可以从两个部分进行分析，首先表示网络带宽的部分考虑了两个节点之间的相互联系，而后面人口乘积开根号相当于单独考虑连接的两个城市的独立信息，所以总体构成是分为每个城市独立的部分以及考虑城市之间相互联系的部分。

考虑条件随机场算法：

$$E = \sum_i E_i + \sum_{i,j} E_{i,j} \quad (3.7)$$

其中 E_i 是 i 节点单独作用产生的损失， $E_{i,j}$ 是将连两个点 i,j 联合作用所产生的损失。这样的算法既考虑了点单独的信息，又考虑了点与点之间的邻域信息。

在我们的网络规划问题中，单独节点的作用规划函数就可以对应于网络价值后半部分的人口计算，而节点之间相互作用产生的价值也就是点与点之间互联的价值也就是价值函数前部分的信道容量。

而现有目标函数虽然具有这样的比较合理的函数形式，但是其中每个部分的设置却有一些不合理的地方。比如在考虑城市单独的价值时，只考虑了城市的人口，却没有考虑城市的发展程度或者生产总值（GDP）的影响因素。所以导致了我们的常理上认为的信息交流的一个中心——上海，只有北京与之相连。而在考虑城市的互联价值时，只考虑了城市与城市之间的距离因素，却没有考虑城市路径上的地形等因素，比如成都和拉萨相连的路径上会有高原地形，导致铺设通信线路难度的增加。

综合以上几点因素，我们查询了网上的相关资料，并考虑城市自身以及城市与城市之间互联的综合因素，提出新的连接价值计算方法：

$$E_{i,j} = \mu C P_{i,j}^{\alpha} G_{i,j}^{\beta} (D_{max} - D_{i,j})^{\gamma} \quad (3.8)$$

其中， C 和 μ 分别表示*i,j*之间相连的通信流量以及地形评分， $D_{i,j}$ 、 $P_{i,j}$ 和 $G_{i,j}$ 分别代表*i,j*城市的距离、人口乘积以及 GDP 乘积， D_{max} 表示可接受最长距离（这里设为 3000）， α, β, γ 皆为可调参数（这里分别设为 0.4, 0.5, 2）。在这里，由于需要综合考虑全国所有的地区，所以我们按照题中所给城市为中心，将全国划分为许多区域，并用这些区域的人口及 GDP 进行计算。

区域	省份	2016年GDP	2016年人口
哈尔滨	辽宁省	22246.9	4378
	吉林省	14776.8	2733
	黑龙江省	15386.09	3799
	总和	52409.79	10910
北京&天津	北京	25669.13	2173
	天津市	17885.39	1562
	河北省	32070.45	7470
	山西省	13050.41	3682
	山东省	68024.49	9947
	总和	156699.87	24834
上海	上海市	28178.65	2420
	江苏省	77388.28	7999
	浙江省	47251.36	5590
	安徽省	24407.62	6196
	总和	177225.91	22205
郑州	河南省	40471.79	9532
武汉	江西省	18499	4592
	湖北省	32665.38	5885
	湖南省	31551.37	6822
	总和	82715.75	17299
西安	内蒙古	18128.1	2520
	陕西省	19399.59	3813
	甘肃省	7200.37	2610
	青海省	2572.49	593
	宁夏	3168.59	675
	总和	50469.14	10211
重庆	重庆市	17740.59	3048
	贵州省	11776.73	3555
	总和	29517.32	6603
成都	四川省	32934.54	8262
昆明	云南省	14788.42	4771
广州&深圳	海南省	4053.2	917
	福建省	28810.58	3874
	广东省	80854.91	10999
	广西省	18317.64	4838
	总和	132036.33	20628
拉萨	西藏	1151.41	331
乌鲁木齐	新疆	9649.7	2398

表格 3-7 主要考虑区域及其辐射区域的人口和 GDP

由于考虑到人口越多所需通信越多，以及 GDP 越高所需通信越多的条件，将两个影响因素相乘的到最终的城市单独参数。同时在距离的考虑上，由于距离越远所需要铺设通信线路的成本越大，所以考虑一个随着距离增加而减少的价值函数。而对于地形参数，我们查阅了中国地形图，发现中国版图大致呈西高东低的三级阶梯形状，分别由昆仑-祁连山脉以及大兴安岭-太行山脉分隔，所以对于跨越分隔山脉的通路进行了降权处理。

而在生成每条边的价值之后，我们设计了特别的整体价值函数，为了避免很多边集中在同样一对节点之间，我们对多个边的权重求和做了衰减，具体函数格式如下：

$$\sum a \frac{E_{i,j}}{\delta + E_{i,j}} = \sum a \frac{\mu CP_{i,j}^\alpha G_{i,j}^\beta (D_{max} - D_{i,j})^\gamma}{\delta + \mu CP_{i,j}^\alpha G_{i,j}^\beta (D_{max} - D_{i,j})^\gamma} \quad (3.9)$$

其中 a , δ 为可调参数， $E_{i,j}$ 同式 (3.8)。

在产生 11 条成树边以后，首先按照选取其他边中最大权重的方法，选取出其他 5 条边。以这 16 条边为初始进行优化算法。我们约定任意两点之间仅会通过一条路径传递信息，并且选取路径中距离最小或总惩罚最小的路径进行传输。在这种假设条件下，我们可以在 16 条边的有环图中使用 Floyd 算法计算出任意两点之间的最短路径，作为这两个城市之间信息传输的路径。有了任意两点之间的传输路径，我们就可以对整个网络中的流量分配使用之前的相同方法进行优化求解，从而得出该结构下的最大价值分配以及所得最大价值。

但是这只是在这种初始网络结构下得到的最优解。而网络结构是有可能变化的。所以我们借鉴模拟退火算法的思想。通常的优化算法需要保证问题的整体结构稳定性，而模拟退火算法可以在问题结构同步更新的情况下找到一个较好的结构以及对应该结构的变量分配。具体来说就是对每次迭代初始的网络结构进行微调，也就是只改变随机一个边的随机一个节点，并且保证整个图的连通性，在新的结构的基础上再次进行相同的优化过程。如果新一次的总价值比上一次有提升，则接受这次的结构更新，否则以概率接受该次的网络结构更新。

而 33 条边的算法也是和 16 条边进行相似的退火优化。

最终由该价值衡量函数及模拟退火算法产生的 11、16、33 条线路图及城市间流量分配如下所示：

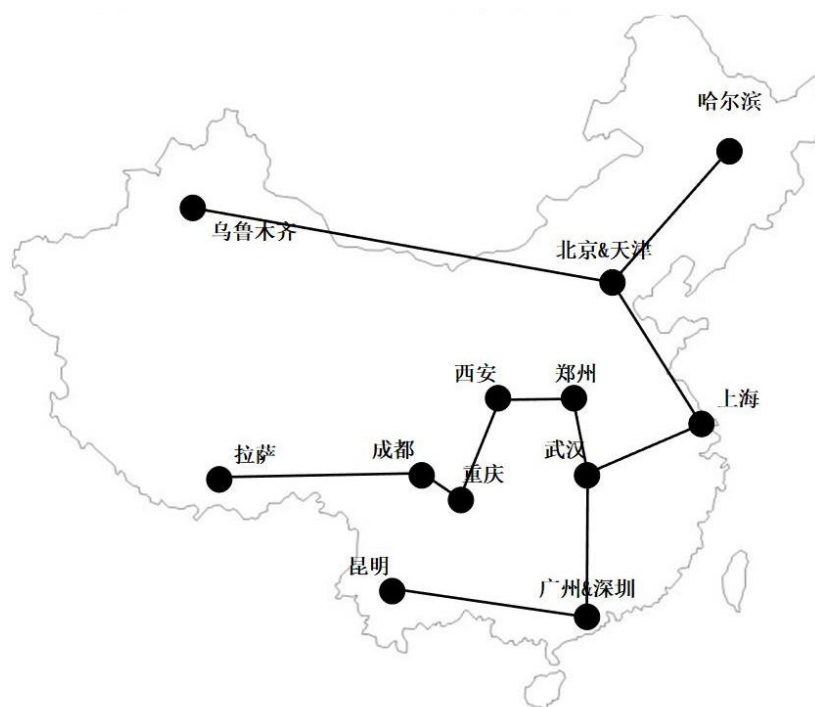


图 3-7 最大生成树连接图

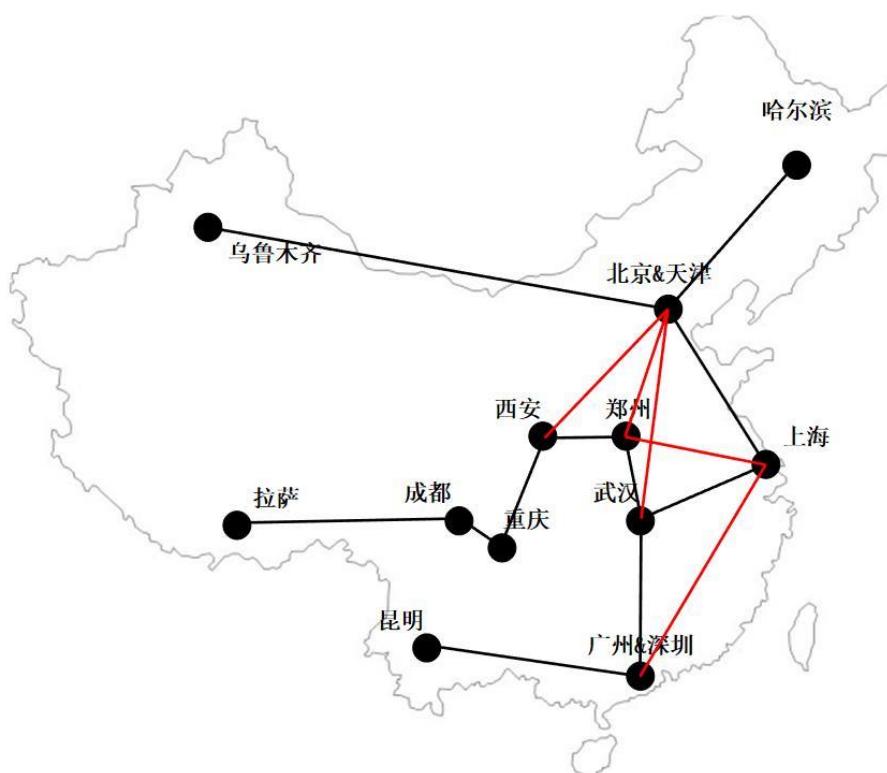


图 3-8 16 连接线网络图

	哈尔滨	北京&天津	上海	郑州	武汉	西安	重庆	成都	昆明	广州&深圳	拉萨	乌鲁木齐
哈尔滨	0	8.279294	0.889463	1.401907	1.117996	1.025855	0.792656	0.792228	0.444787	1.047813	0.104255	0.103745
北京&天津		0	1.012223	1.52177	1.274261	1.114924	0.945346	0.869306	0.567683	0.760943	0.109093	0.103744
上海			0	2.775181	3.744123	1.83089	1.021138	0.934156	0.585675	2.999412	0.103988	0.103751
郑州				0	4.250288	3.697272	1.808591	1.554385	0.869168	2.712232	0.215043	0.722318
武汉					0	2.57111	1.381799	1.214058	0.72981	3.831481	0.931356	0.10375
西安						0	2.660674	2.181639	0.996793	1.77275	0.759572	0.103761
重庆							0	4.456281	1.857985	1.100329	3.537199	0.140931
成都								0	2.058474	0.996383	1.029181	0.103755
昆明									0	0.573871	1.007268	6.308487
广州&深圳										0	0.101037	0.103749
拉萨											0	0.102009
乌鲁木齐												0

表格 3-8 16 条边时城市间的通信流量分配

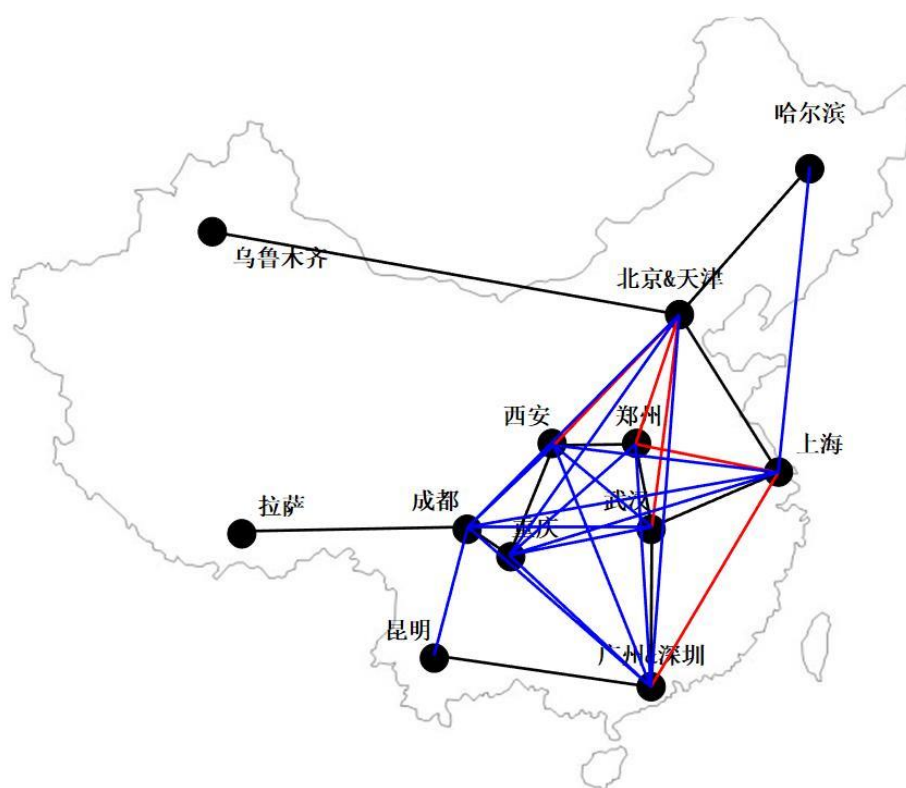


图 3-9 33 连接线网络图

	哈尔滨	北京&天津	上海	郑州	武汉	西安	重庆	成都	昆明	广州&深圳	拉萨	乌鲁木齐
哈尔滨	0	10.30692	1.025589	1.234962	0.937102	0.619025	1.234898	0.213142	0.110964	0.10904	0.103977	0.104376
北京&天津		0	1.190057	3.328885	2.016439	0.759135	0.489969	0.4727	0.110097	1.525227	0.171398	0.243018
上海			0	2.592854	3.807213	0.257609	1.185289	0.29099	0.753482	3.407972	1.384537	0.104409
郑州				0	3.549948	0.946273	0.400068	0.354648	3.777403	4.008982	0.209245	6.634904
武汉					0	2.162651	0.973653	2.441182	0.111093	4.408158	2.738794	0.100452
西安						0	5.064953	2.639205	5.315896	1.672815	0.151377	0.170786
重庆							0	2.670415	4.657311	0.376105	2.543406	0.146197
成都								0	0.757394	0.104308	0.313312	0.104391
昆明									0	0.111274	0.110137	0.184949
广州&深圳										0	0.17171	0.10441
拉萨											0	0.102107
乌鲁木齐												0

表格 3-9 33 条边时城市间的通信流量分配

可以看出，在新的价值计算方法中，各条线路规划变得更加合理。华南地区以武汉和上海为中心，华北地区以北京为中心，并且对于西南地区以及新疆的规划也都更加合理，同时也仅有较少的保持连通所必要的跨地形阶梯连接。如果未来考虑更多因素，我们也可以将线路与线路之间的一些混合关系加入价值计算体系当中，来得到更加优化的网络规划结果。

4 问题三：改善星座图

4.1 符号说明

符号	意义
P_s	信号的平均功率
P_n	噪声的平均功率
SNR	信号噪声功率比（信噪比）
N	调制方式的阶数，代表对应的星座图上的符号个数
K	星座图上的点 $K \in \{1, \dots, N\}$
\mathbf{s}_k	理想的接受到的符号向量
\mathbf{n}_k	噪声向量
\mathbf{r}_k	实际接收的符号向量
$PDF(k)$	星座点 k 的概率密度函数
D_k	星座点 k 的判决区域
d_{kj}	星座点 k 判决为星座点 j
$err(d_{kj})$	星座点 k 判决为星座点 j 时的错误 bit 数
p_k	星座点 k 的概率
Ω	信息熵，定义为 $\Omega = -\sum_1^N p_k \log_2(p_k)$

4.2 模型的建立

1. 目标函数

最小化容限点 SNR，使其小于 8QAM 的 SNR 容限；

2. 约束条件

a) 信号功率归一化约束：

$$P_s = \frac{1}{N} \sum_{k=1}^N |s_k|^2 = 1$$

b) 信息熵约束：

$$\Omega = -\sum_1^N p_k \log_2(p_k) = 3$$

c) 信噪比等式关系

$$SNR = \frac{P_s}{P_n}$$

d) 高斯噪声方差约束

$$\sigma^2 = P_n = 1/SNR = 10^{\frac{-SNR(dB)}{10}}$$

e) 高斯分布约束

$$PDF_{QAM} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x - k_x)^2 + (y - k_y)^2}{2\sigma^2}\right)$$

f) 错误比特概率

$$E_{bit} = \sum_{k=1}^N P(k) \times \sum_{j=1}^N err(d_{kj}) \times \iint_{D_{k_j}} PDF(k) dx dy$$

g) 容限前误码率约束

$$BER = \frac{E_{bit}}{\Omega} = 0.02$$

3. 模型变量

- a) 星座点的概率 $p_k, k \in K$
- b) 星座点的坐标 $s_k, k \in K$

4.3 问题分析

根据约束条件：利用 16QAM 来传输 3bit 信息熵的信息。在信息论中，熵是接收的每条消息中包含的信息的平均量，又被称为信息熵、信源熵、平均自信息量。如果 16 个星座点是等概率分布的，那么得到的信息熵是最大的，为 4bit。同时 8QAM 的调制分布等概率分布才会得的最大的信息熵 3bit。因此，需要提出的优化星座图的星座点数 N 必须满足 $8 < N \leq 16$ ，且必须是非等概率分布。

当变换星座图上星座点的概率时，如果某个星座点的概率 $p_k = 0$ 时，就表示其在星座图上不再有效，此时 N 小于 16。考虑有效范围内不同的 N 的值，当 N 比较大时，星座图内的星座点的个数较多，星座点中间的判决空间就较易产生重合，星座点判决出错的概率就较大。因此 N 需要尽量较小。而如果 $N=8$ ，则需要等概率的星座图分布，就和 8QAM 没有区别了。因此本文着力尝试构造 9QAM 的星座图。受到文献[1]的启发，我们尝试了基于霍夫曼编码来构造 9 个不等概率星座点的方式。

为了使星座点偏移出错的概率较小，同问题一中一样，我们规定每个点同最邻近点的距离相同为 d 。

4.4 问题求解

Step1: 根据霍夫曼编码方式构造一种 9QAM 的星座点的概率分布；

Step2: 选取 16QAM 上适合的点对 9QAM 进行格雷码编码；

Step3: 依据功率归一化公式求得具体的星座点坐标；

Step4: 依据星座点的位置划分判决区域；

Step5: 根据二维高斯分布的概率密度函数计算每一种判决 d_{ij} 的概率分布；

Step6: 结合星座点的概率 $P(k_i)$ 和概率密度函数得到错误比特数期望；

Step7: 计算 SNR-BER 关系图；

Step8: 使用二分法不断逼近求解 $BER = 0.02$ 的 SNR 容限点；

首先，根据约束条件(a),(b)，我们通过穷举 2 的负幂次方的方式寻找到了的一组有效的概率分布的解，如下表所示

p_k	个数
$\frac{1}{4}$	1
$\frac{1}{8}$	4
$\frac{1}{16}$	4

为了保证所有的九个点之间有格雷码编码的特性，本文采取原 16QAM 星座图上左下角的 9 个点的编码方式作为该 9QAM 的编码方式。其余的点概率为 0 舍弃。同时保持原来 16QAM 的星座点的位置关系，保持两两最邻近点之间的等距关系，得到一个 3*3 的星座点阵。

考虑功率归一化条件，我们将概率较大的点分配在离原点较近的坐标上以最大化星座点间的距离 d 。因此我们得到如下星座点分布；

星座点	概率	类型
0000	1/4	I
0100	1/8	II
0001		
1000		
0010		
1010	1/16	III
0110		
0101		
1001		

表格 4-1 9QAM 的星座点分布

将星座点表示在复平面上，9QAM 的星座图以及判决区域如图 4-1。9-QAM 的星座图符合格雷码编码，因此相邻的星座点之间只有 1bit 的错误。

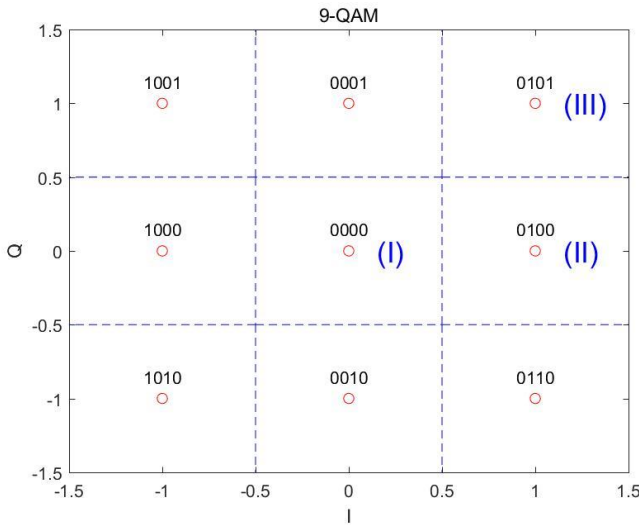


图 4-1 9-QAM 的星座图及判决区域

下面求出该星座图的 SNR 容限点，根据功率归一化的约束，求的 $d = 1\sqrt{mW}$ 对于 I 类点，以星座点 (0000) 为例子，其概率密度函数为：

$$PDF_{9QAM-I} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

其错误 bit 数的期望为：

$$E_{bit}^I = 4 \times \iint_{D_{0100}} PDF_{9QAM-I} ds + 2 \times \iint_{D_{0101}} PDF_{9QAM-I} ds$$

对于其他类型的点采用相同的思路，求得其错误 bit 数的期望 $E_{bit}^{II}, E_{bit}^{III}$

BER 的计算公式为：

$$BER_{9QAM} = \frac{\frac{1}{4} \times E_{bit}^I + \frac{1}{8} \times 4 \times E_{bit}^{II} + \frac{1}{16} \times 4 \times E_{bit}^{II}}{\Omega}$$

在 Matlab 中根据公式计算出 BER 和 SNR 的曲线如图 2-4 所示。由于 BER-SNR 的函数是单调的，通过二分法逼近得到精确的容限点 $SNR_{9QAM}(BER = 0.02) = 8.7321$

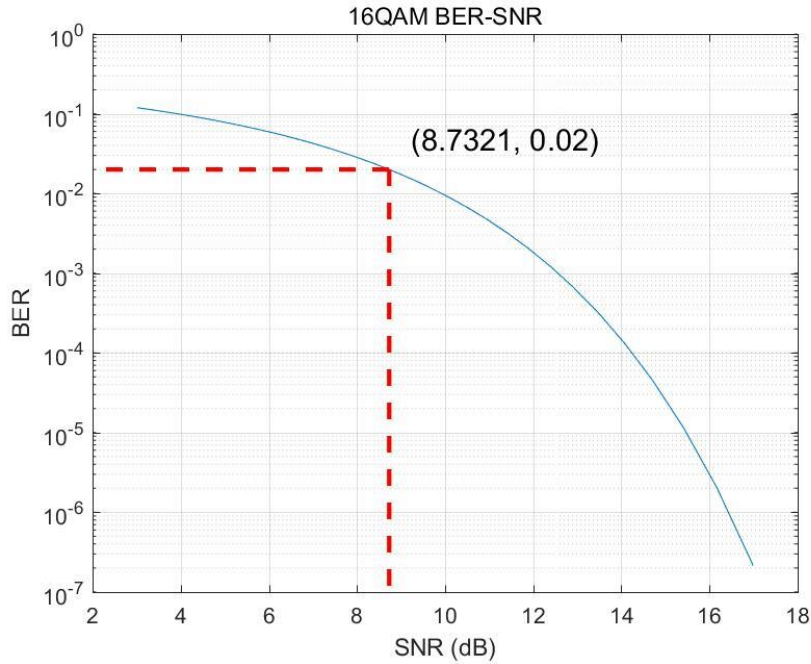


图 4-2 16-QAM BER-SNR 关系曲线

4.5 结果

通过对题目的分析和模型求解，我们求得了一个信息熵为 3bit 的 9QAM 非等概率编码方式。其 SNR 容限点为 8.7321。

图 4-3 给出了其与题目中设定的 8QAM 的调制格式的性能曲线的比较结果。

可以看到在 $BER = 0.02$ 时，本文提出的 9QAM 调制格式能将 SNR 容限点降低 1.6 dB。

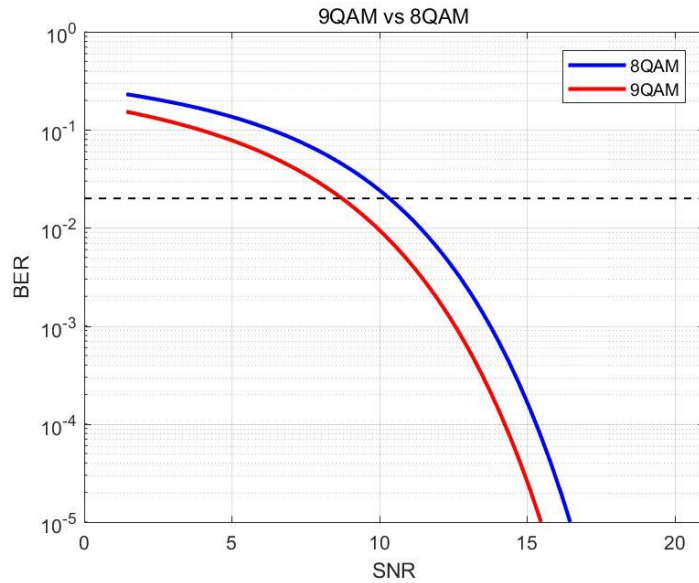


图 4-3 优化的 9QAM 与 8QAM 的 BER-SNR 性能曲线对比

5 模型讨论与评价

5.1 模型的优点

- 1、为解决问题一(1)，我们采用了二维正态分布模型，通过面积积分来计算纠错前误码率。这是对星座图调制误码率的一个最为准确的估计模型。
- 2、问题二，对网络模型进行适当优化，减少问题复杂度，可以很方便的找出最大生成树。在树已经确定的情况下对整体网络流量进行分配，保证达到该条件下的全局最优解。
- 3、问题三中得到的 9QAM 调制模型成功的找到一个优于题设 8QAM 的调制模式，其将 8QAM 的 SNR 容限点降低了 1.6Db

5.2 模型的缺点及改进方向

- 1、问题二中，需要对树的结构进行提前确认，有时可能会错失属于另外树结构下的全局最优解。
- 2、问题三中得到的 9QAM 限制了星座点的点数以及其概率分布，仅仅是一个局部最优解，而非全局最优解。

6 参考文献

- [1] 石悦 邱雪松, Y. Shi, S. QIU Xue, Y. GUO Shao 和 F. Qi, 《基于改进遗传算法的电力光传输网规划方法, Optimal planning of optical transmission network using improved genetic algorithm》, 通信学报
- [2] A. Eshraghniaye Jahromi 和 Z. Besharati Rad, 《Optimal topological design of power communication networks using genetic algorithm》, Scientia Iranica
- [3] 陈启美 李博等 《通信原理教程》[M].南京: 南京大学出版社。2008
- [4] Cho K, Yoon D. On the general BER expression of one- and two-dimensional amplitude modulations[J]. IEEE Trans Commun, 2002, 50(7):1074-1080.
- [5] Liu T, Lin C, Djordjevic I B. Advanced GF(32) nonbinary LDPC coded modulation with non-uniform 9-QAM outperforming star 8-QAM[J]. Optics Express, 2016, 24(13):13866.
- [6] 中华人民共和国国家统计局, 全国各项分省年度统计数据 <http://data.stats.gov.cn/easyquery.htm?cn=E0103, 20180917>
- [7] 司守奎 孙兆亮 《数学建模算法与应用》 国防工业出版社. 2017